Spontaneous self-assembly of engineered armadillo repeat protein fragments into a folded structure

Watson, Randall P ; Christen, Martin T ; Ewald, Christina ; Bumbak, Fabian ; Reichen, Christian ; Mihajlovic, Maja ; Schmidt, Elena ; Güntert, Peter ; Caflisch, Amedeo ; Plückthun, Andreas ; Zerbe, Oliver

Abstract: Repeat proteins are built of modules, each of which constitutes a structural motif. We have investigated whether fragments of a designed consensus armadillo repeat protein (ArmRP) recognize each other. We examined a split ArmRP consisting of an N-capping repeat (denoted Y), three internal repeats (M), and a C-capping repeat (A). We demonstrate that the C-terminal MA fragment adopts a fold similar to the corresponding part of the entire protein. In contrast, the N-terminal YM2 fragment constitutes a molten globule. The two fragments form a 1:1 YM2:MA complex with a nanomolar dissociation constant essentially identical to the crystal structure of the continuous YM3A protein. Molecular dynamics simulations show that the complex is structurally stable over a 1 s timescale and reveal the importance of hydrophobic contacts across the interface. We propose that the existence of a stable complex recapitulates possible intermediates in the early evolution of these repeat proteins.

DOI: https://doi.org/10.1016/j.str.2014.05.002

Supplementary Information to the Paper Titled "Spontaneous Self-Assembly of Engineered Armadillo Repeat Protein Fragments into a Folded Structure"

by R. P. Watson *et al.*

Content

Tables

Figures

Cloning, related to "Experimental Procedures: Section Cloning, Expression and Purification"

Oligonucleotides were purchased from Microsynth AG (Balgach, Switzerland).

Table S1: Primers used for the expression of the various fragments, related to "Experimental Procedures: Section Cloning, Expression and Purification"

Name	Sequence 5'-3' direction	Purpose
LIC_M_for	GAAAATTTATATTTTCAGGGGAACGAACAAATCCAAGCTGTTATCGATGC	MA, M_2A
LIC_C_rev	AGATGAGAGTAAGGCTATCATTAGTGGGACTGCAGCTTCTCCAGAGC	MA, M_2A
LIC_N_for	GAAAATTTATATTTTCAGGGGGAACTGCCGCAGATGACCCAGCAGCTGAACTCC	YM_2
LIC_M_rev	AGATGAGAGTAAGGCTATCATTAACCAGAAGCGATGTTAGACAGAGCCCACAGAGC	YM_2
102_dTrp_pLIC_for	GAAAATTTATATTTTCAGGGGAAAGCAATTTTCGTACTG	YM
103_dTrp_3C_rev	GGGCCCCTGGAACAGCACTTCCAGCTG	YM
104_YM_3C_for	GTGCTGTTCCAGGGGCCCGGGGAACTGCCGCAGATG	YM
105_YM-pLIC_rev	AGATGAGAGTAAGGCTATCATTAACCAGAAGCGAT	YM
53_YMRx_for	TAATGAGGTACCCCGGGTCGACCTGCAGCC	YMR, YMRRR
54_YMRRR_rev	CAGTTCAGCGATGTTAGTCAGAGCGTCCAG	YMRRR
56_YMR_rev	AGCGAAAGCGATGTTGTTCAGAGCGATAAG	YMR

Protein Expression, related to the Section "Self-Assembly of a Split Consensus Armadillo Repeat Protein"

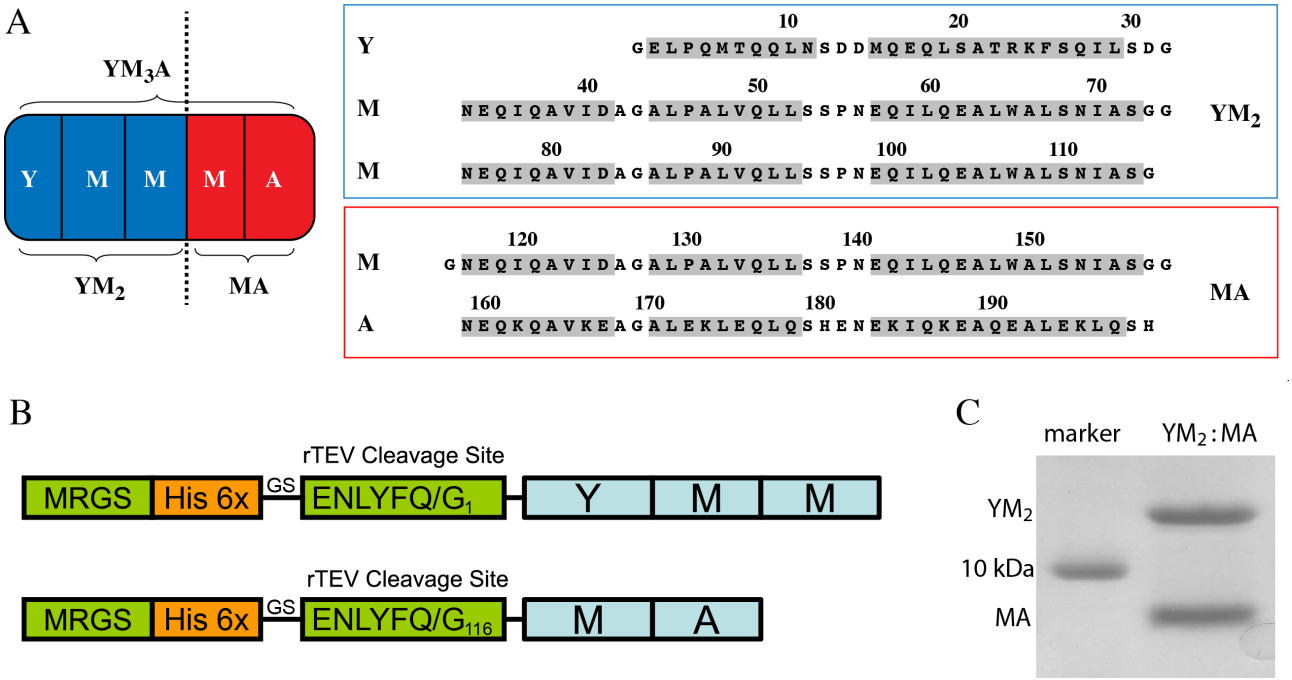

Figure S1 related to the Section "Self-Assembly of a Split Consensus Armadillo Repeat Protein": A, Amino acid sequences of the two main repeat protein fragments YM_2 and MA investigated in this study. **B**, Schematic overview of expression products YM_2 (top left) and MA (bottom left). **C**, Coomassie-stained 15% SDS-PAGE analysis of the rTEV treated, Ni-column purified products before further purification by SEC.

Table S2: Theoretical molecular weights of unlabeled fragments, related to "Experimental Procedures: Section Cloning, Expression and Purification"

Construct w/o isotopic labeling	MW [kDa]
YM_2 w/o His-tag	12.2
MA w/o His-tag	9.1
YM_2 : MA complex both w/o His-tag	21.3
YM w/o His-tag	8.0
YMR incl. His-tag	13.8
YMRRR incl. His-tag	22.6
M_2A w/o His-tag	13.4

Size Exclusion Chromatography and Multi-Angle Light Scattering (MALS) Analysis

Table S3 related to Figure 3: Comparison of molecular weights based on the sequence, SEC, and MALS analysis.

Construct	MW_{calc} [kDa]	MW_{SEC}[a] [kDa]	Ratio MW_{SEC}/MW_{calc}	MW_{MALS}[b] [kDa]
YM_2	12.2	33.8	2.8	13.3 ± 1.5
MA	9.1	24.9	2.7	8.8 ± 0.5
YM_2:MA	21.3	35.3	1.7	18.4 ± 1.4
YM_3A	21.3	35.3	1.7	19.4 ± 0.5

a. normalized value from all SEC measurement

b. average of 2 measurements per concentration at 100, 50 and 25 µM protein.

Circular Dichroism Spectra, related to Section "Heteronuclear NMR Demonstrates that the Fragments Interact Specifically"

Both YM_2 and MA displayed α-helical (208 and 222 nm) characteristics, however, YM_2 appears considerably less structured (Figure S2A). YM_3A and the YM_2:MA complex also show α-helical characteristics.

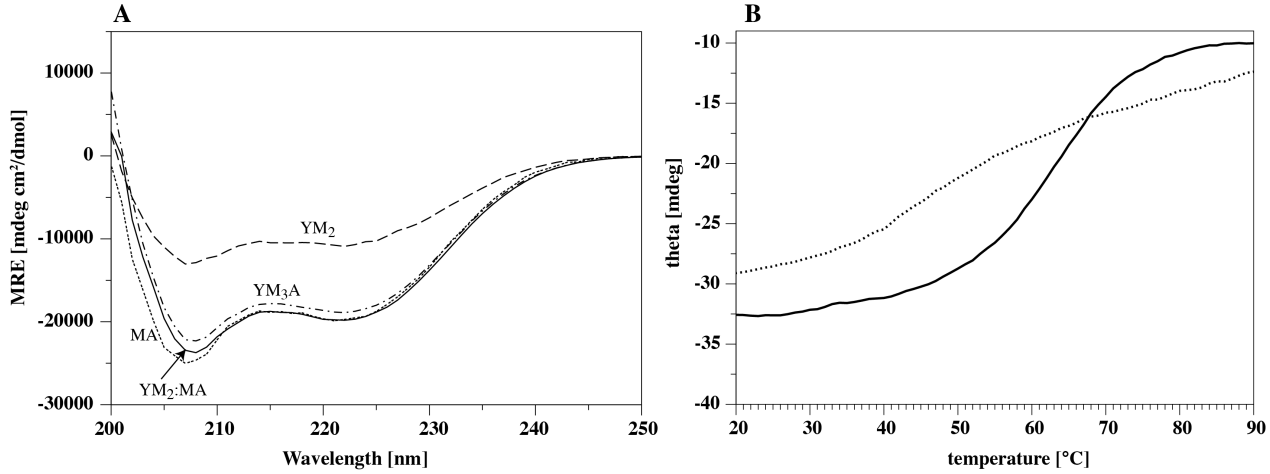

Figure S2, related to Section "Heteronuclear NMR Demonstrates that the Fragments Interact Specifically": **A**, CD spectra of YM_3A (dash-dotted line) and YM_2:MA (1 equiv. each, solid line), YM_2 (dashed line) and MA (dotted line). **B**, Thermal denaturation observed at 220 nm for YM_2 (dotted line) and MA (solid line).

The melting point of MA and YM_2 was recorded on a JASCO J-715 (JASCO PFD425S Peltier-controlled).

Final measurement concentrations for the melting curves were: 2 µM protein in 5 mM sodium phosphate buffer, pH 7.4, 15 mM NaCl, 0.2 % glycerol. Path length 1 cm, slit width 1 nm, integration time 0.125 s, wave length 220 nm, heating rate 1°C min^{-1}. The melting point of MA was found to be 62°C, whereas an exact melting temperature for YM_2 could not easily be determined due to the flat character of the curve without a true transition point or plateau (Figure S2B).

Other split sites, related to Section "Heteronuclear NMR Demonstrates that the Fragments Interact Specifically"

We have also probed for other split sites in the YM_3A and YMRRRMA proteins. The sequences of the protein fragments are depicted in Table S4:

Table S4 related to Section "Heteronuclear NMR Demonstrates that the Fragments Interact Specifically": Amino acid sequences of fragments used to investigate other split sites. Randomized positions in "R" repeats are indicated in red. Unintentional point mutations acquired during the selection process are indicated in yellow.

YM	1 6 11 16 21 26 31 36 41 46 51 56 61 66 71 GPGEL PQMTQ QLNSD DMQEQ LSATR KFSQI LSDGN EQIQA VIDAG ALPAL VQLLS SPNEQ ILQEA LWALS NIASG
YMR	1 6 11 16 21 26 31 36 41 46 51 56 61 MRGSHHHHHH GSELP QMTQQ LNSDD MQEQL SATVK FRQIL SRDGN EQIQA VIDAG ALPAL VQLLS SPNEQ ILQEA 66 71 76 81 86 91 96 101 106 111 116 LWALS NIASG GNEQT QAVID AGALP ALVQL LSSPN EQILQ YALIA LNNIA FA
YMRRR	1 6 11 16 21 26 31 36 41 46 51 56 61 MRGSHHHHHH GSELP QMTQQ LNSDD MQEQL SATVK FRQIL SRDGN EQIQA VIDAG ALPAL VQLLS SPNEQ ILQEA 66 71 76 81 86 91 96 101 106 111 116 121 126 131 136 LWALS NIASG GNEQT QAVID AGALP ALVQL LSSPN EQILQ YALIA LNNIA FAGNE QTQAV IDAGAL PALVQ LLSSP 141 146 151 156 161 166 171 176 181 186 191 196 NGQIL QETLW ALTNI AMEGN EQQQA VIDAG ALPAL VQLLS SPNEQ ILQYA LDALT NIAEL
M2A	1 6 11 16 21 26 31 36 41 46 51 56 61 66 71 GNEQI QAVID AGALP ALVQL LSSPN EQILQ EALWA LSNIA SGGNE QIQAV IDAGA LPALV QLLSS PNEQI LQEAL 76 81 86 91 96 101 106 111 116 121 126 WALSN IASGG NEQKQ AVKEA GALEK LEQLQ SHENE KIQKE AQEAL EKLQS H

Spectra for ^{15}N-labeled N-terminal fragments when complexed to the complementary unlabeled C-terminal fragments are depicted in Figure S3:

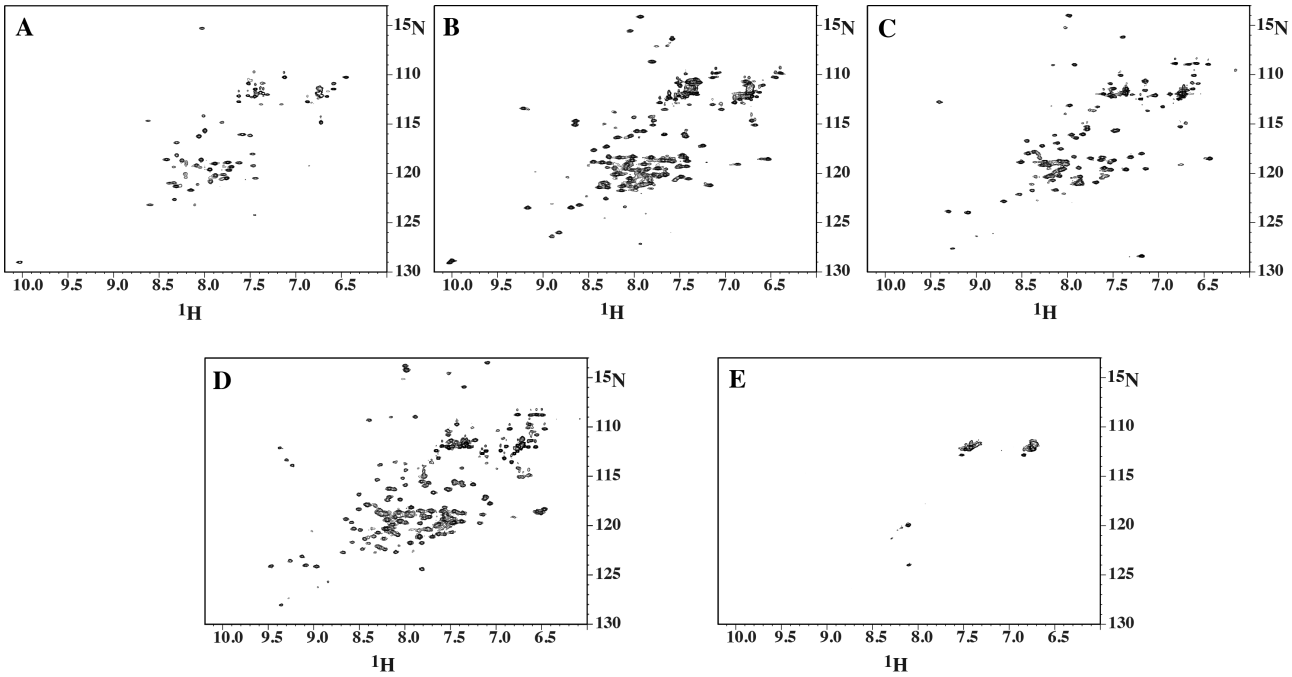

Figure S3 related to Section "Heteronuclear NMR Demonstrates that the Fragments Interact Specifically": 700 MHz [^{15}N,^{1}H]-HSQC spectra of **A**, YM:M_2A (YM 250 µM); **B**, YM_2:MA (YM_2 500 µM); **C**, YMR:M_2A (YMR 250 µM) and **D**, YMRRR:MA. (YMRRR 250 µM). Only the N-terminal fragment is ^{15}N-labeled, excess of unlabeled fragments 1.5 fold. Uncomplexed YMRRR is shown for reference in **E**.

NMR Spectroscopy related to "Experimental Procedures, Section NMR Spectroscopy, Assignments and Structure Calculation"

The following experimental spectra were collected for both the free MA fragment and the YM_2:MA complex: 2D [^{15}N,^{1}H]-HSQC (Müller, 1979, Bodenhausen and Ruben, 1980) and constant-time [^{13}C,^{1}H]-HSQC (Vuister and Bax, 1992); 3D HNCO (Marion et al., 1989), HN(CA)CO (Clubb et al., 1992), HNCACB (Wittekind and Mueller, 1993), HN(CO)CACB (Grzesiek and Bax, 1992), HCCH-TOCSY (Bax et al., 1990, Olejniczak et al., 1992, Kay et al., 1993) and HN(CO)CCCH (Grzesiek and Bax, 1992). Additionally, we collected 4D HCCH-TOCSY and 4D HCCH-NOESY spectra of YM_2 in presence of unlabeled MA to aid in the sequential assignment of the highly repetitive complex. Both of the latter experiments were recorded with sparse sampling (about 5 % of data points). All 2D and 3D experiments utilized TPPI-States for quadrature detection in indirect dimensions (Marion et al., 1989), and gradient-based coherence selection (echo-antiecho) in combination with sensitivity enhancement schemes for experiments that detect amide protons (Kay et al., 1992).

All spectral data collected for the YM_2:MA complex were later converted with CCPN FormatConverter (Vranken et al., 2005) to Azara or USCF format to provide compatibility with the CCPN Analysis 2.3.1 (Vranken et al., 2005) and Sparky 3.115 (Goddard and Kneller) softwares, respectively, which were used to refine the assignments.

Sequential assignment related to Section "Structures of the Fragments in the Complex Closely Mimic the Structure of the Covalently Linked Full-length Armadillo Protein"

The YM_2 and MA Armadillo fragments were assigned individually based on spectra from standard triple-resonance experiments to annotate the ^{15}N,^{1}H and ^{13}C,^{1}H correlation maps derived from the [^{15}N,^{1}H]-HSQC and constant-time [^{13}C,^{1}H]-HSQC spectra, respectively (Figures S4 and S5). For the uncomplexed MA fragment, we were able to annotate 77.5 % of all the backbone amides in the construct, which rose to 92.8 % for backbone and 85.9 % for sidechain resonances when ignoring the missing first 13 amino acids. When complexed with YM_2, these 13 residues within the MA fragment assume a predominantly helical conformation and we were able to assign 92.1 % of the backbone amide resonances and 85.6 of all protons.

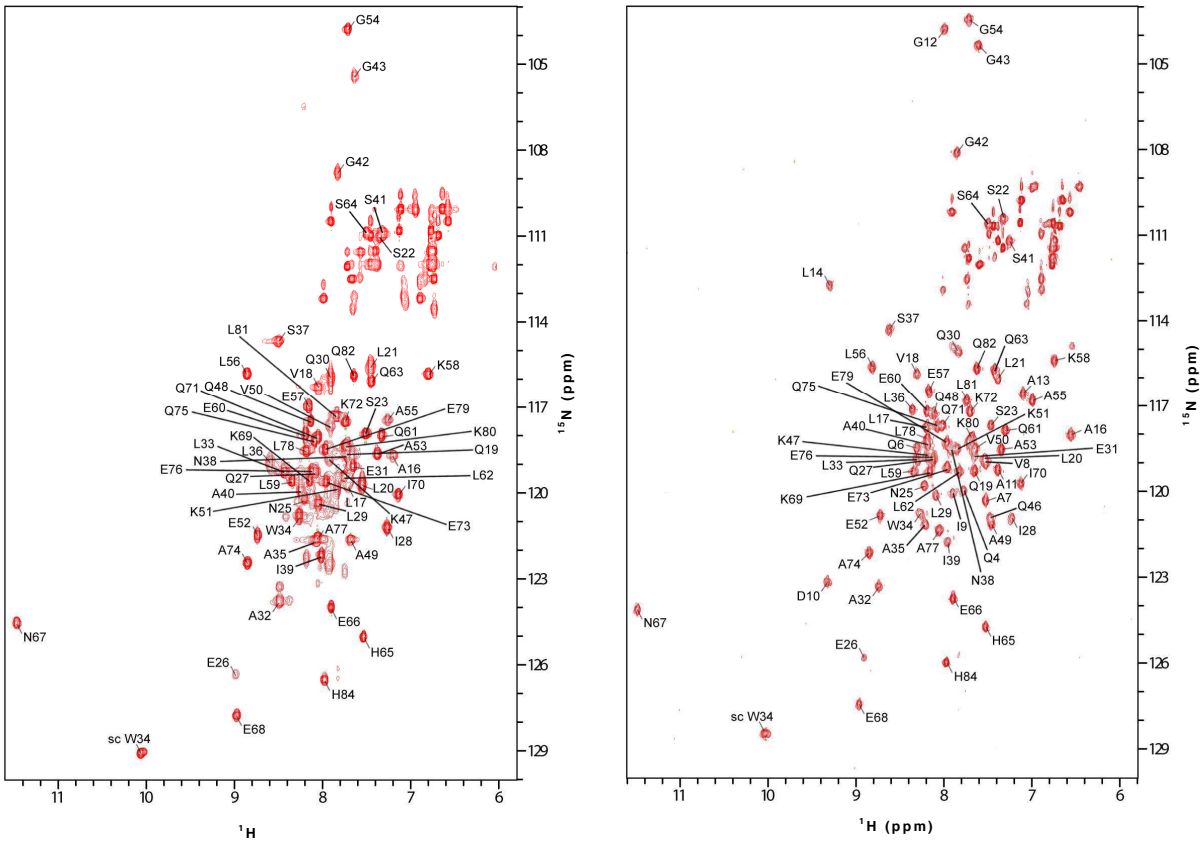

Figure S4 related to Figure 1: 600 MHz [^{15}N,^{1}H]-HSQC spectrum of MA in uncomplexed (left) or complexed with 1.2 equiv. YM_2 (right) forms recorded at 32°C, 0.75 mM in PBS_{150} pH 7.4, 2 % glycerol, 10 % D_2O, 1 mM TMSP, 0.02 % NaN_3.

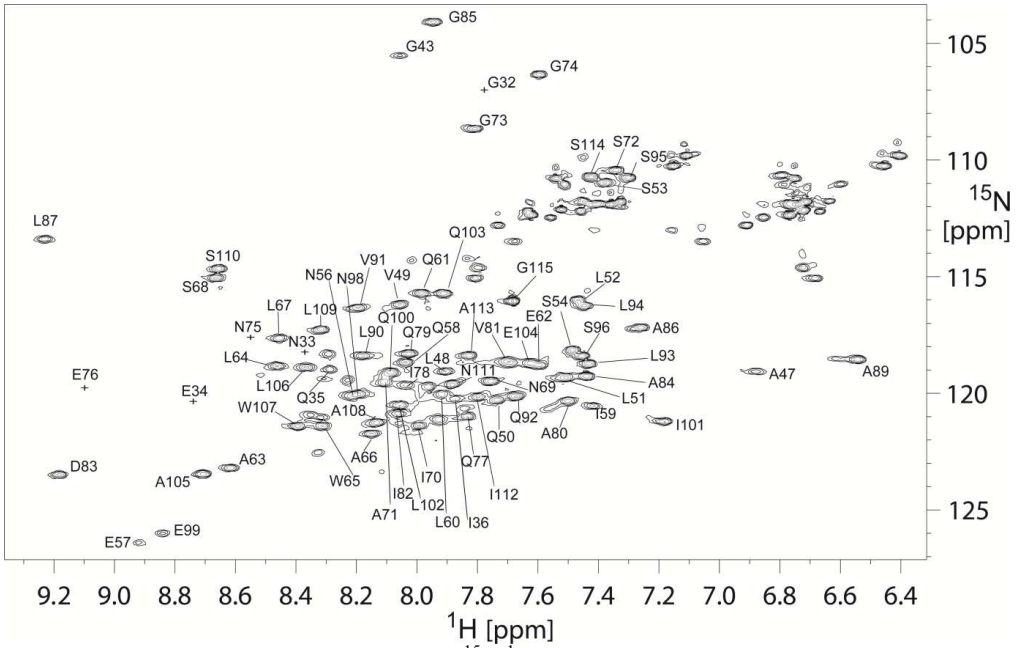

Figure S5 related to Figure 1: 700 MHz [^{15}N,^{1}H]-HSQC spectrum of YM_2 complexed with 1.5 equiv. of MA, recorded at 32°C, 1 mM in PBS_{150} pH 7.4, 2 % glycerol, 10 % D_2O, 1 mM TMSP, 0.02 % NaN_3.

Structure calculation and validation of the YM_2:MA complex related to the Section "Structures of the Fragments in the Complex Closely Mimic the Structure of the Covalently Linked Full-length Armadillo Protein"

Initially, separate YM_2 and MA chemical shift data were used to automatically assign NOE crosspeaks from ^{13}C- and ^{15}N-resolved 3D HSQC-NOESY spectra (τ_{mix} = 75 ms), and determine separate preliminary structures for the two complexed fragments (not shown) using the protocol applied for uncomplexed MA outlined in the main paper. In a second step, interfacial H–H distances < 5 Å were extracted from the crystallographic structure of the corresponding single-chain Armadillo construct (PDB entry 4DBA). Used in conjunction with the preliminary structures of the fragments, we thus generated a synthetic NOE peaklist to help guide the assignment process, yielding 66 manually identified interfacial distance restraints. For the final structure calculation, all available data was collated and curated in CCPN Analysis 2.3.1 (Vranken et al., 2005) before being used as input for structure calculation as follows: a total of eight 3D NOESY spectra (4 for each fragment, encompassing both filtered and unfiltered ^{13}C-edited and ^{15}N-edited data) were linked to individual chemical shift lists that were filtered to include only the theoretically observable resonances for each spectrum type. The interfacial distance restraints identified manually in step two above were artificially loosened by 1 Å and added to the structure calculation in order to partially constrain the complex while allowing for local rearrangement. Using all these restraints, augmented by the TALOS-N derived dihedral restraints for YM_2 and MA (Figure S6), a structural ensemble was calculated for the entire complex with UNIO'10 (Guerry and Herrmann, 2012). The calculated bundle was restrained by 1916 unambiguous distances (404 long-range, out of which 77 are located at the interface) and 279 dihedral torsion angles (Table 2).

Inspection of the structural bundle of the complex initially obtained from UNIO revealed that its high target function (~33 $Å^2$) is almost exclusively due to steric clashes at the repeating Leu-Pro dipeptide (i.e. at positions 45-46, 87-88 and 129-130). To address this problem and simultaneously maximize sampling of the conformational space, the closed-ring Pro residues at positions 46, 88 and 130 were replaced by their open-ring equivalents (CYANA residue code PROO) and corresponding intra-residue constraints added to allow for pyrrolidine ring puckering flexibility during the simulated annealing step. These new constraints, combined with those previously identified by UNIO, were leveraged to calculate 1000 new structures via CYANA 3.96 (Güntert, 2004). The 100 lowest total energy conformers yielded a bundle with an improved target function of ~9.4 $Å^2$. For consistency, this refinement protocol was applied to the uncomplexed MA fragment as well. In the final iteration, the conformational ensembles for both structures were subjected to refinement in explicit TIP3P water using the parallhdg5.3 parameters implemented in the nmr_waterrefine extension (Linge et al., 2003, Nabuurs et al., 2004) to XPLOR-NIH 2.35.

Each conformer in the YM_2:MA bundle was analyzed for secondary structure using the software STRIDE (Heinig and Frishman, 2004) followed by visual inspection. A plot of restraints per residue reveals that some sites along the primary structure are more restrained than others (Figure S7, panel A). Comparing the number of inter-residual NOEs detected against the sequential RMSD values yields Pearson product-moment correlation coefficients (ρ) of $\rho_{NOE(inter),\ RMSD(bb)}$ = -0.448 and $\rho_{NOE(inter),\ RMSD(hv)}$ = -0.458, confirming the existence of a moderate linear correlation between the number of restraints and the local precision of the calculated structure. Residues in the vicinity of helical elements at stretches 60 – 92, 100 – 114, 120 – 156, 171 – 179 and 186 – 196 are well

localized and essentially coincide with those in the x-ray structure. Conversely, the greatest divergence between structures in the bundle is observed for the N-terminal cap and residues 159 and 160 (Figure S7, panel B). The local precision of the NMR bundle is mirrored (Figure S7, panel C) in the profile of root-mean-square-fluctuations (RMSF) obtained from molecular dynamic (MD) simulations of the lowest energy conformer (*vide infra*). The helical regions, notably H3, are substantially more rigid than the connecting loops and the RMSD/RMSF profiles of the individual modules remain similar over all three repeats.

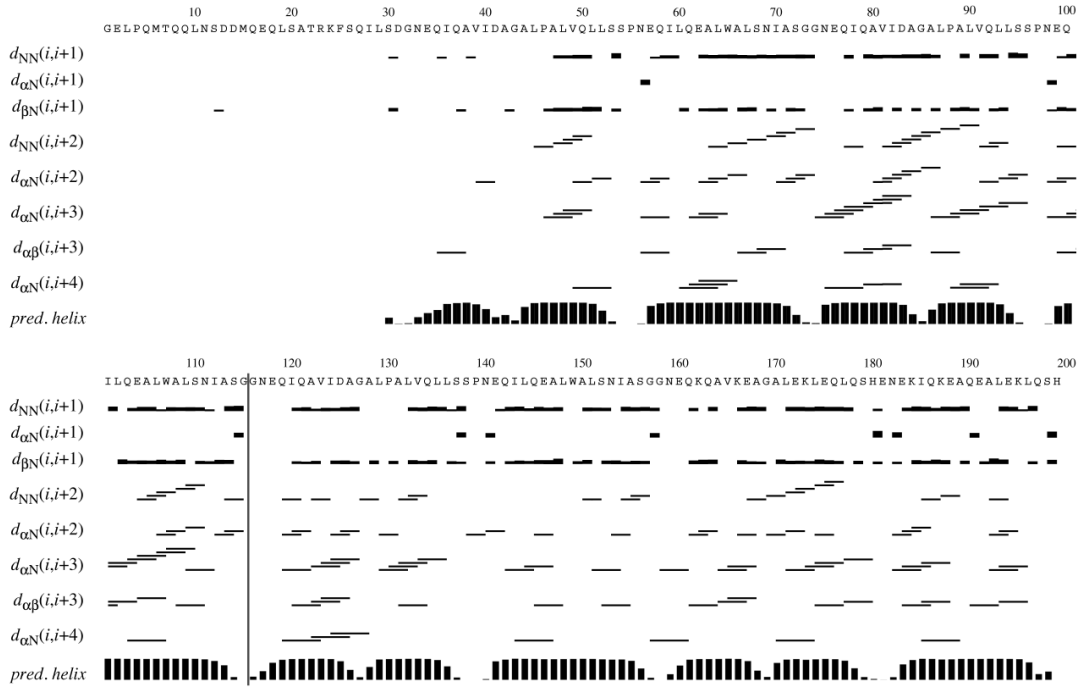

Figure S6 related to Figure 5: NMR experimental short-range distance constraints for the YM_2:MA complex. Constraints are shown *versus* amino acid sequence (top) and helical structure elements predicted by TALOS-N (bottom). Thickness of the horizontal bars reflects relative intensities (weak, medium, strong) of the sequential and short-range NOEs. The locus of the separation between fragments YM_2 and MA is indicated by a vertical line.

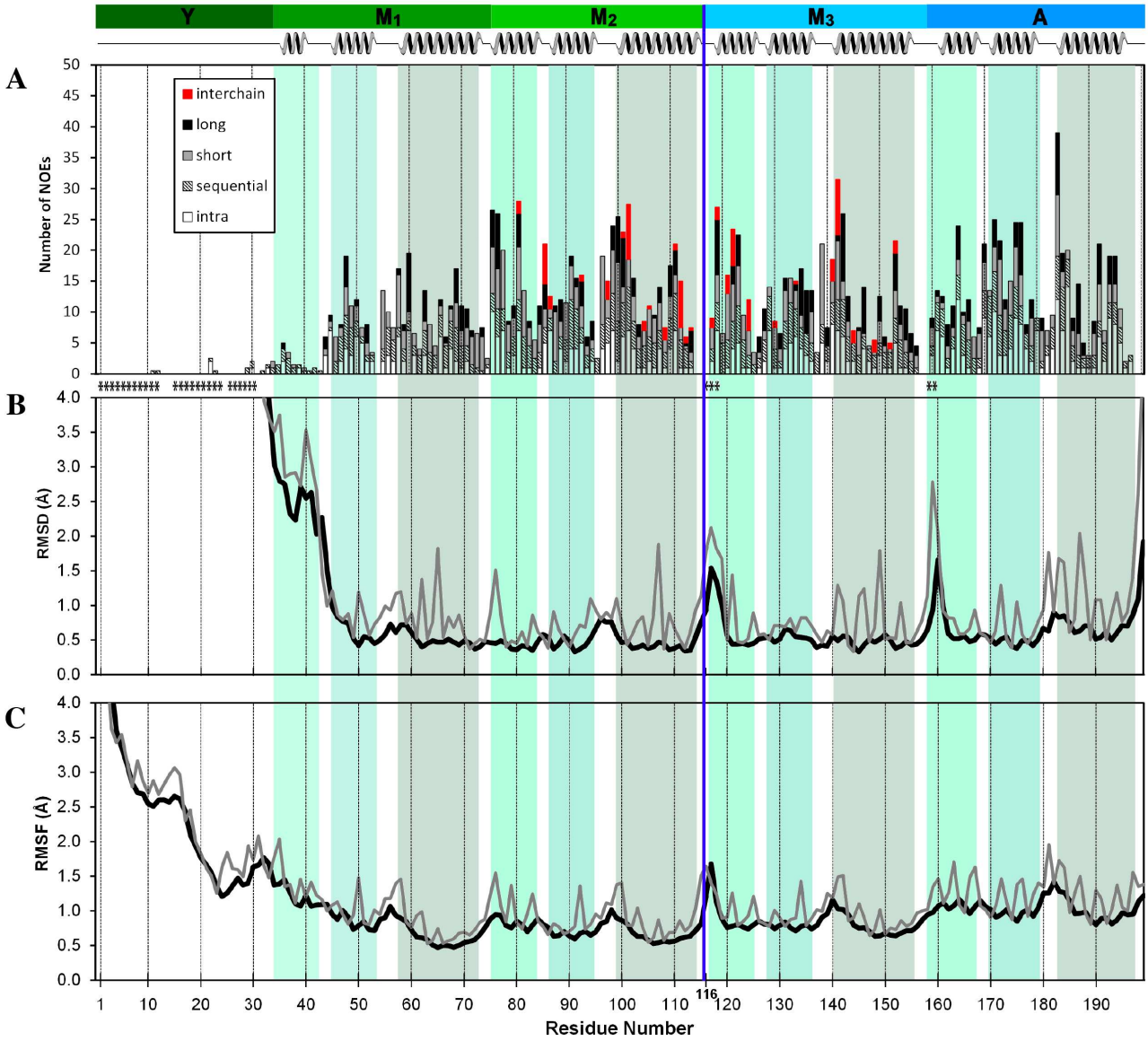

Figure S7 related to Figure 6: Number of NMR distance constraints *versus* precision of the lowest-energy structure of the bundle of the armadillo complex structures. Loci of the modular repeats and secondary structure are schematically shown at the top; the separation between the two fragments preceding residue 116 (labeled) is indicated by a vertical blue line. **A**, Number of NOEs versus residue number. NOEs are: intra-residue (□, $i \rightarrow i$), sequential (□, $|i - j| = 1$), short range (■, $1 < |i - j| < 5$), long range (■, $|i - j| \geq 5$) and inter-chain (■, between YM_2 and MA). Residues with unassigned/undetected Cα resonances are indicated (*). **B**, Heavy atom average RMSD to the mean of the protein backbone (dark trace) and protein backbone plus side chains (light trace), calculated for a bundle of 20 computed NMR structures. Note that residues 159 and 160 could not be assigned, which leads to a local lack of constraints. **C**, Sequence profile of the root-mean-square-fluctuation (RMSF) of the backbone (black) and heavy (gray) atoms during an MD simulation started from the lowest-energy NMR conformer. RMSF values were calculated on the 250 2-ns segments between 500 ns and 1000 ns, and then averaged. The six remaining MD runs (four more from the lowest-energy NMR conformer and two from the crystal structure) show essentially identical RMSF profiles except for smaller displacements of the Y cap (residues 1-34) in the two runs started from the crystal structure. For reference, the location of the helices predicted from the crystal structure is indicated by color-shaded boxes.

MD Simulations

We have performed five MD simulations starting from the coordinates of the NMR structure of the YM_2:MA complex (Table S5). Three of these MD runs were initiated from the lowest-energy NMR conformer of the complex, each starting with a different seed for the random assignment of the initial velocities (NMR1-NMR3). The fact that no resonances of residues 1-30 could be assigned for the NMR structure determination due to signal broadening indicates that those residues are not simply disordered, but rather undergo conformational change in the intermediate exchange régime. To account for this, we have prepared two more NMR conformers, 4DBA-NMR and 4DB6-NMR, in which residues 1-30 were modelled from crystal structures with PDB entries 4DBA and 4DB6, respectively. The former entry, which contains the identical amino acid sequence, displays an unexpected domain swap in the N-cap; the latter is a monomer that displays the expected N-cap conformation (*i.e.* intramolecular association with the rest of the protein) but with a different amino acid sequence in the N-cap. This N-cap conformation was mutated *in silico* and grafted onto the remainder of the NMR solution complex structure to ensure consistency in the protein sequence. MD calculations from all these NMR-derived structures were performed to study a potential influence of the N-cap on the stability of the complex. Despite the differences in the modeled N-cap, the multiple runs from the X-ray structure give a very similar result, which indicates statistical robustness. Two further MD runs were started from the crystal structure 4DBA (labelled "xtal") and from an artificially split x-ray conformer (labelled "split-xtal", see main text of the paper).

Table S5 related to Section "MD Simulations": Summary of the MD simulations

Trajectory name	Starting structure
NMR1, NMR2, NMR3[a]	Lowest energy NMR conformer of the YM_2:MA complex
4DBA-NMR	Residues 1-30 of the crystal structure (PDB: 4DBA, domain swap) grafted onto the lowest energy NMR conformer of the YM_2:MA complex
4DB6-NMR	Residues 1-30 of the crystal structure (PDB: 4DB6) grafted onto the lowest energy NMR conformer of the YM2:MA complex
xtal	Crystal structure (PDB: 4DBA)
split-xtal	Crystal structure (PDB: 4DBA) with the "hydrolyzed" amide bond between G115 and G116, i.e. G115-$COO^{-}\cdots H_3N^{+}$-G116

[a] Initial velocities for independent runs NMR1, NMR2, and NMR3 were assigned using different seeds for the random number generator.

In the following we analyze the motion and energetics along the MD trajectories, and in particular the RMSD from the initial structure (Figure S8), the location of secondary structure (Figure S9), and interaction energies between neighboring repeats (Figure S10). The time series of RMSD show that all MD simulations behave qualitatively similarly irrespective of the starting structure (Figure S8 and Figure 8 in the main text).

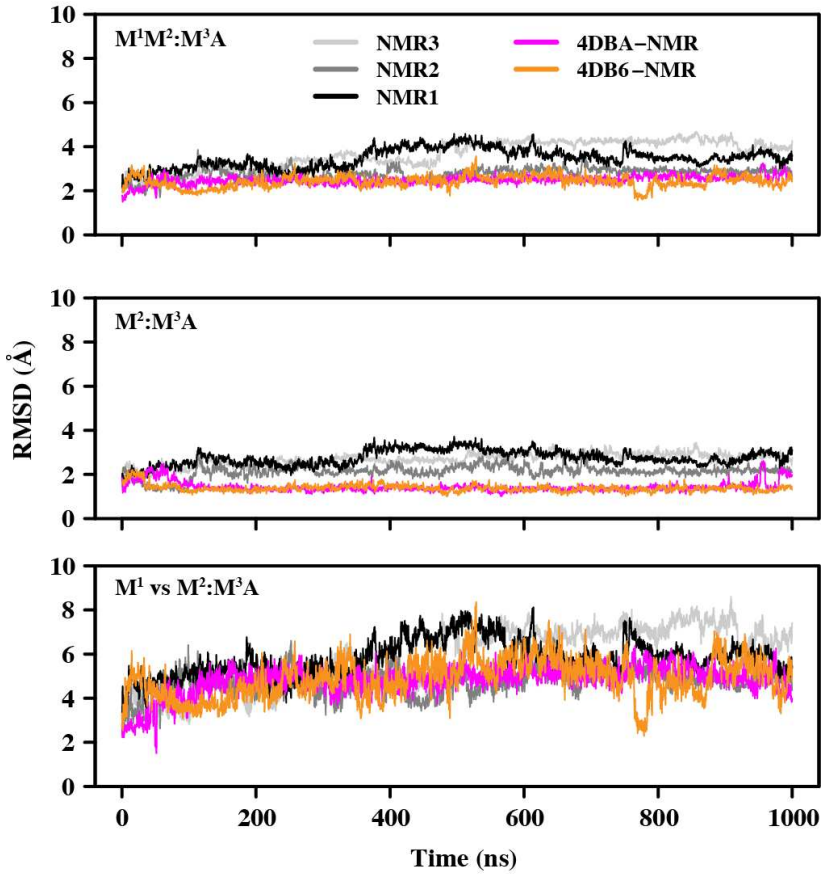

Figure S8 related to Figure 8: Structural stability of the YM_2:MA complex in the MD simulations (see Table S5 for their descriptions). The time series of the root mean square deviation (RMSD) from the X-ray structure (PDB code 4DBA) were calculated for the Ca atoms of repeats $M^1M^2M^3A$ (upper panel) (i.e. not considering the N-cap), M^2M^3A (middle panel), and the first M-repeat of the N-terminal fragment YM^1M^2 upon fitting repeats M^2M^3A (lower panel) to the crystal structure 4DBA. For other simulations see Figure 8 of the main text.

Presence and location of secondary structure, *i.e.* DSSP annotations (Kabsch and Sander, 1983), were calculated using the CAMPARI software (Vitalis and Pappu, 2009) and are shown in Figure S9.

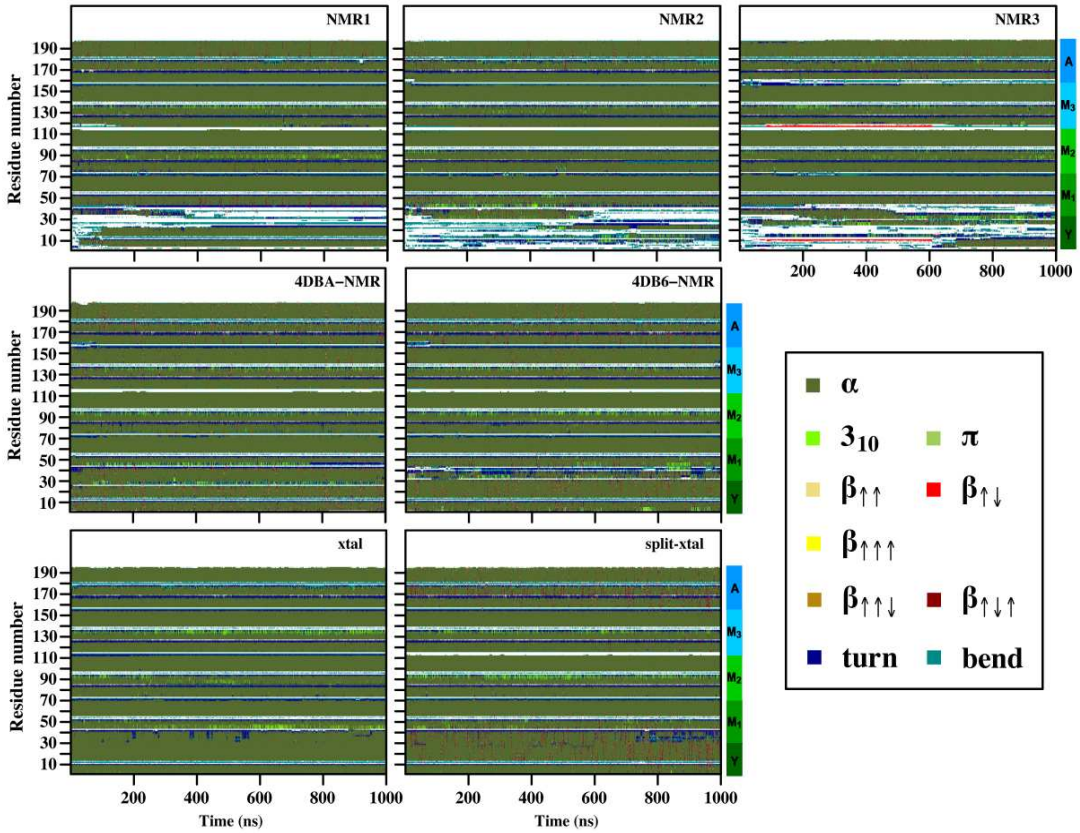

Figure S9 related to Figure 8. Time series of secondary structure along the MD simulations for the MD runs NMR1-3, 4DBA-NMR, 4DB6-NMR, xtal and split-xtal. Note that secondary structure in the domain-swapped N-cap is largely retained during the simulation.

A time length of 1 µs is not sufficient for the equilibration of the disordered N-terminal segment (residues 1-34), since different conformations are sampled depending on the starting conformation. The time series demonstrate that there is only partial formation of helical structure in runs NMR1-3, whereas secondary structure in the others runs remains intact in the N-cap, as it was present in the starting conformation.

To further shed light on the structural stability, we analyzed the interaction energies between neighboring repeats and decomposed the total interaction energy into Coulomb and van der Waals contributions (Figure S10). The time series of interaction energies reveals that the van der Waals energy is more favorable than the Coulomb energy for M^1M^2 and M^2M^3, while both energies are of comparable magnitude for M^3A. It is important to note that the temporal evolution of individual energy terms provides support for statistical convergence except for the large fluctuations in the interactions between Y and M^1, which is due to the large displacements of Y.

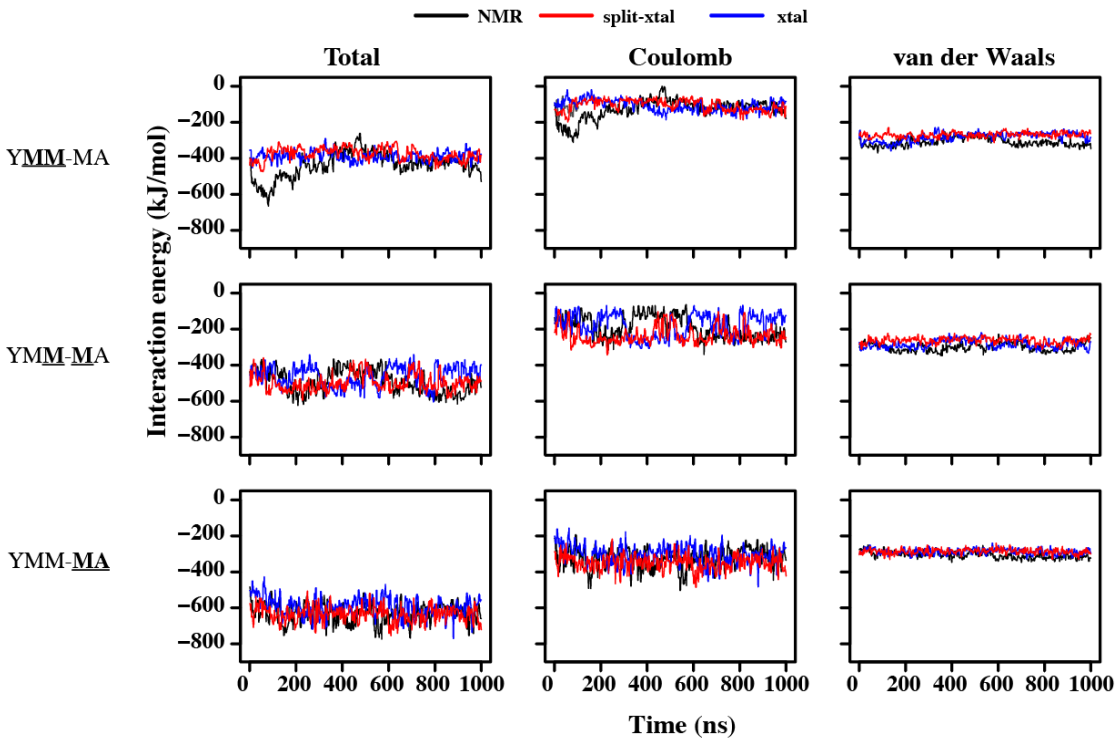

Figure S10 related to Section "MD Simulations": Time series of interaction energy between the neighboring repeats of YM_3A along the MD simulations from run NMR1 or from simulation of the molecule as found in the crystal structure (either as entire protein (xtal) or split protein (split-xtal). The total interaction energy (left panels) is the sum of Coulomb energy (middle panels) and van der Waals energy (right panels). The neighboring repeats for which energies are calculated are labelled in bold and underlined on the left. A similar behavior was observed for the four remaining simulations that were started from the NMR structure.

References:

1. Bodenhausen, G., and Ruben, D.J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. Chem. Phys. Lett. *69*, 185–189.
2. Müller, L. (1979) Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. J. Am. Chem. Soc. *101*, 4481–4484.
3. Vuister, G.W., and Bax, A. (1992) Resolution enhancement and spectral editing of uniformly C-13-enriched proteins by homonuclear broad-band C-13 decoupling. J. Magn. Reson. *98*, 428–435.
4. Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M., and Clore, G.M. (1989) Overcoming the Overlap Problem in the Assignment of ^{1}H NMR Spectra of Larger Proteins by Use of Three-Dimensional Heteronuclear ^{1}H-^{15}N Hartmann-Hahn Multiple-Quantum Coherence and Nuclear Overhauser-Multiple Quantum Coherence Spectroscopy: Application to Interleukin 1β. Biochemistry *28*, 6150–6156.
5. Clubb, R.T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue ^{1}HN, ^{15}N, and ^{13}C'chemical shifts in ^{15}C-labelled proteins. J. Magn. Reson. *97*, 213–217.
6. Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha-

and beta-carbon resonances in proteins. J. Magn. Reson. *101*, 201–205.
7. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. J. Am. Chem. Soc. *114*, 6291–6293.
8. Kay, L.E., Xu, G.Y., Singer, A.U., Muhandiram, D.R., and Forman-Kay, J.D. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins. J. Magn. Reson. Ser. B *101*, 333–337.
9. Olejniczak, E.T., Xu, R.X., and Fesik, S.W. (1992) A 4D HCCH-TOCSY experiment for assigning the side chain ^{1}H and ^{13}C resonances of proteins. J. Biomol. NMR *2*, 655–659.
10. Bax, A., Clore, G.M., Driscoll, P.C., Gronenborn, A.M., Ikura, M., and Kay, L.E. (1990) Practical aspects of proton-carbon-carbon-proton three-dimensional correlation spectroscopy of ^{13}C labelled proteins. J. Magn. Reson. *87*, 620–627.
11. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) Rapid recording of 2D NMR spectra without phase cycling. application to the study of hydrogen exchange in proteins. J. Magn. Reson. *85*, 393–399.
12. Kay, L.E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J. Am. Chem. Soc, *114*, 10663–10665.
13. Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687–696.
14. Goddard, T.D., and Kneller, D.G. SPARKY 3.
15. Guerry, P., and Herrmann, T. (2012) Comprehensive automation for NMR structure determination of proteins. Methods Mol. Biol. *831*, 429–451.
16. Güntert, P. (2004) Automated NMR structure calculation with CYANA. Methods Mol. Biol. *278*, 353–378.
17. Linge, J.P., Williams, M.A., Spronk, C.A., Bonvin, A.M., and Nilges, M. (2003) Refinement of protein structures in explicit solvent. Proteins *50*, 496–506.
18. Nabuurs, S.B., Nederveen, A.J., Vranken, W., Doreleijers, J.F., Bonvin, A.M., Vuister, G.W., Vriend, G., and Spronk, C.A. (2004) DRESS: a database of refined solution NMR structures. Proteins *55*, 483–486.
19. Heinig, M., and Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. *32*, W500–W502.
20. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577–2637.
21. Vitalis, A., and Pappu, R.V. (2009) Methods for Monte Carlo simulations of biomacromolecules. Annu. Rep. Comput. Chem. *5*, 49–76.